

Geoinformatik – aktuelle Forschungsfragen

Sester, Monika

Veröffentlicht in:
Jahrbuch 2007 der Braunschweigischen
Wissenschaftlichen Gesellschaft, S.93-103



J. Cramer Verlag, Braunschweig

Geoinformatik – aktuelle Forschungsfragen*

MONIKA SESTER

Institut für Kartographie und Geoinformatik, Leibniz Universität Hannover
Appelstraße 9 a, D- 30167 Hannover

Zusammenfassung

Geoinformatik beschäftigt sich mit der Erfassung, Verwaltung und Analyse raumbezogener Daten und Informationen. Damit hat sie viele Berührungspunkte zu Nachbardisziplinen wie Informatik, Geodäsie, Kartographie, Geographie aber auch Mathematik. Im Beitrag werden schlaglichtartig aktuelle Forschungsfragen beleuchtet und anhand von Arbeiten am Institut für Kartographie und Geoinformatik dargestellt.

1. Einführung und Übersicht

Geodaten werden mit verschiedenen Sensoren erfasst. Sie stehen in unterschiedlichsten Formen und Datentypen bereit; im Wesentlichen kann man Bilddaten und Vektordaten unterscheiden. Die Vektordaten sind meist semantisch annotiert, d.h. es lassen sich beispielsweise Straßen, Gebäude oder Biotopgrenzen unterscheiden. Die Geodaten werden über standardisierte Schnittstellen bereitgestellt und finden sich bekanntermaßen in analogen und digitalen Karten, in Fahrzeugnavigationssystemen, aber auch in Mobiltelefonen, die ihre Nutzer über sogenannte Location Based Services über das nächste italienische Restaurant oder das nahe Blumengeschäft informieren können. Sehr bekannt sind heute virtuelle Globen wie Google Earth, das seine Nutzer durch z.T. hoch aufgelöste Luft- und Satellitenbilder, aber auch Zusatzinformation in Form von überlagerten Straßen oder sonstigen Points-of-Interest informiert.

Die Fülle, Vielfalt und steigende Verfügbarkeit von Geodaten stellt Chancen und Herausforderungen für die Forschung dar. Zum einen liegen sie in der Handhabung der Informationsflut, zum anderen in der geeigneten Integration und Kombination für neue Fragestellungen, z.B. für Anwendungen in der Navigation.

* Kurzfassung des am 09.11.2007 vor der Plenarversammlung der Braunschweigischen Wissenschaftlichen Gesellschaft gehaltenen Vortrags.

2. Informationsflut

a. Dateninterpretation

In der Geodäsie werden traditionell Messsensoren eingesetzt, mit der die 2D- bzw. 3D-Realität erfasst und rekonstruiert werden kann. Standen in der Vergangenheit spezielle Messgeräte im Vordergrund, mit denen dediziert eine Messung ausgewählter, bedeutungsvoller Punkte durchgeführt wurden, kommen in jüngerer Zeit Sensoren hinzu, mit denen Massendaten erfasst werden, in denen sich das Objekt indirekt abbildet. Dies ist der Fall bei Bilddaten aber auch bei den neuen Laserscanning-Daten, in denen die Geometrie von Objekten in Form von 3D-Masspunkten erfasst wird (siehe Abbildung 1). Kennzeichen für diese Art von Daten ist, dass die Objekte darin implizit gegeben sind (in der Anordnung der 3D-Punkte) und eine Weiterverarbeitung in der Regel eine Interpretation der Punktmenge erfordert. So können beispielsweise Ebenen in den Punktwolken gesucht werden, die dann die einzelnen Fassadenelemente des Gebäudes darstellen können (siehe Abbildung 2). Diese Ebenen müssen anschließend zu einer konsistenten 3D-Gebäudegeometrie zusammengesetzt werden. Hierfür wurde im Kontext von luftgestützten Laserdaten ein Verfahren entwickelt, welches in der Lage ist, vollautomatisch 3D-Stadtmodelle abzuleiten (Brenner, 2005).



Abb. 1: 3D-Laserscan des Welfenschlosses der Leibniz Universität Hannover.



Abb. 2: Segmentierung der 3D-Punktwolke in einzelne Ebenen.

b. Raumbezogene Suchmaschinen

Suchmaschinen sind heute für viele ein nützliches Instrument im Alltag geworden. Raumbezogene Informationen spielen dort eine zunehmende Rolle. Wenn es darum geht, lokale Anfragen zu stellen, z.B. die Suche nach einem Zahnarzt oder einer Apotheke, spielt die räumliche Nähe natürlich eine essentielle Rolle. Auch Anfragen nach touristischen Zielen beinhalten meist räumliche Aspekte, die durchaus von heutigen Systemen teilweise schon gelöst werden. Anfragen der Form „Hotel in Hannover“ können heute bereits gelöst werden und die Suchmaschinen liefern sehr gute Hinweise auf relevante Webseiten. Probleme stellen sich, wenn es darum geht, weitere räumliche Konstrukte zu nutzen, die der Mensch in seiner Sprache oft verwendet, etwa „Hotel in der Nähe des Flughafens“ oder gar Negationen wie „Hotel möglichst außerhalb des Kneipenviertels“. Diese Anfragen lassen sich mit gängigen Suchmaschinen, die auf dem Prinzip des Vergleichs mit Schlüsselbegriffen arbeiten, nicht ohne Weiteres lösen. Ein Beispiel ist die Anfrage „Hotel im Norden von Hannover“, welches alle Webseiten zurückliefert, welche die Schlüsselbegriffe „Hotel“, „Norden“ und „Hannover“ enthalten. Dies bringt neben einer richtigen Antwort, in der das Hotel explizit auf seiner Webseite vermerkt hat, dass es sich im Norden Hannover befindet, auch ein Hinweis auf eine Reiseagentur namens „Buche den Norden“ (siehe Abbildung 3).

Im Rahmen des EU-Projektes SPIRIT wurden am ikg zu raumbezogenen Suchmaschinen geforscht. Es wurden Möglichkeiten geschaffen, einige räumliche Relationen zu modellieren und damit für Abfragen nutzbar zu machen. Abbildung 4 zeigt das Ergebnis einer Anfrage nach „Schloss in der Nähe von Hannover“.

Ein weiteres Problem für die Integration von Geodaten in Suchmaschinen ist, dass diese in der Regel durch ihre Geometrie repräsentiert werden und ihre Bedeutung erst durch ihre Form und die räumlichen Relationen gewinnen. Bei-

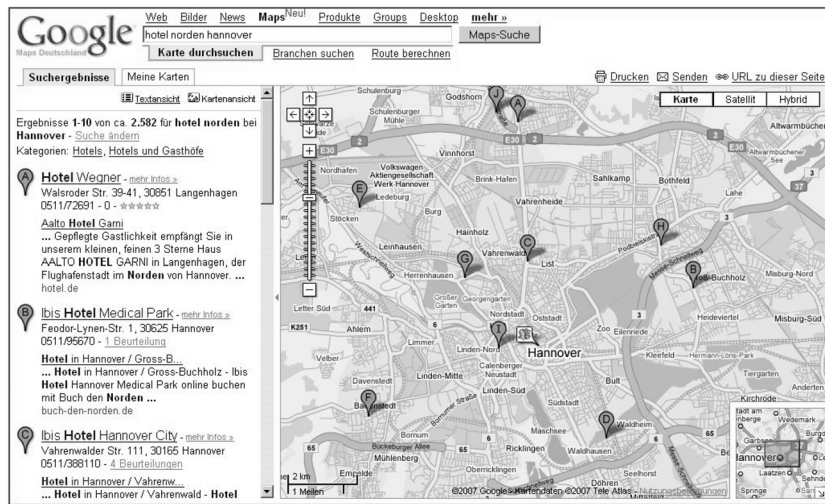


Abb. 3: Suchanfrage nach „Hotels im Norden von Hannover“.

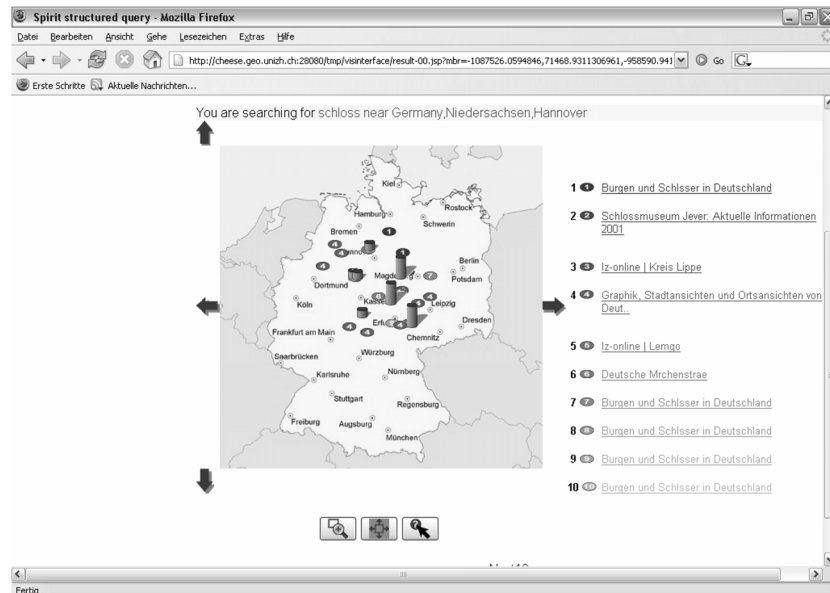


Abb. 4: Suchanfrage an räumliche Suchmaschine SPIRIT: Schloss near Hannover.

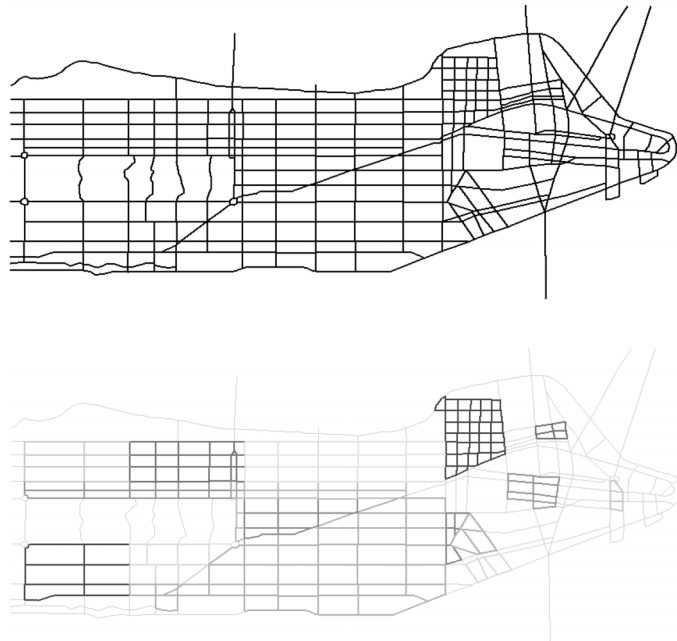


Abbildung 5: Straßennetz (oben), sowie automatisch extrahierte Gitterstrukturen (unten).

spielsweise fällt es uns Menschen nicht sehr schwer, in einem Datensatz bestimmte Objekte zu erkennen, etwa in Abbildung 5 ein Straßennetz, ohne dass diese explizit als solche bezeichnet sind. Um jedoch gezielt nach Geodaten zu suchen, müssen diese bekannt und benannt sein.

Mittels Methoden der Algorithmischen Geometrie und des Spatial Data Minings lassen sich solche Strukturen automatisch finden und somit für Suchanfragen zugänglich machen (HEINZLE et al., 2006). HEINZLE (2007) entwickelte Methoden mit denen sich Straßenstrukturen in beliebigen Datensätzen automatisch extrahieren lassen, wie z.B. Gitter, Ringstrukturen oder auch sternförmig zusammenlaufende Straßen (siehe Abb. 5).

c. Datenabstraktion –Generalisierung

Die Beherrschung und Kommunikation der Datenflut erfordert eine Abstraktion der Information. Dies wird auf graphischem Wege in der Kartographie traditionell durch sogenannte Generalisierungsverfahren durchgeführt, welche zu Karten in unterschiedlichen Maßstäben führen. Auf diese Weise wird es möglich,

eine hierarchische grob-zu-fein-Annäherung an die Information durchzuführen. An der Automation dieser Prozesse wird derzeit international geforscht. Eine Aufgabenstellung ist dabei die sogenannte Typifizierung, d.h. die Reduktion einer gegebenen Objektmenge unter Beibehaltung der räumlichen Verteilung. Diese wird beispielsweise nötig, wenn es darum geht, Gebäude in kleineren Maßstäben darzustellen. Ein Blick in Karten der Maßstäbe 1:50.000 bzw. 1:100.000 zeigt dass dort die Gebäude stark vergrößert und lediglich als Symbole, meist Quadrate, dargestellt sind, da sie in ihrer wahren Größe dargestellt zu klein wären, um noch sichtbar zu sein. Die Reduktion der Anzahl der Objekte muss dabei derart erfolgen, dass deren räumliche Verteilung beibehalten bleibt. Im folgenden wird ein automatisches Verfahren zur Typifizierung vorgestellt.

2.1. Typifizierung auf Basis von Kohonen Merkmalskarten

Der am ikg entwickelte Ansatz basiert auf Kohonen Merkmalskarten (SESTER, 2005), einem Verfahren aus dem Bereich der Neuronalen Netze (KOHONEN 1982). Gegeben ist ein sogenannter Eingaberaum E der Dimension m mit Reizen x , sowie ein Kartenraum A der Dimension d mit Neuronen, wobei d üblicherweise ein- oder zweidimensional ist. Jedes Neuron im Kartenraum wird durch ein Tupel $U=(w,p)$ beschrieben, d.h. einem Gewicht w aus E und einer Position p aus A . Aus der Position p können die Nachbarn des Neurons bestimmt werden. Das Gewicht w gibt an, auf welchen Reiz das Neuron am stärksten reagiert. Häufig werden diese Positionen in äquidistanten, diskreten Gittern vorgegeben. Die a priori Gewichtung erfolgt zufällig, wenn über die Struktur der Karte keine Information vorliegt. Andernfalls kann diese Information einfließen.

Der Algorithmus lässt sich wie folgt beschreiben:

- Stimuluswahl: Aus der Menge der Reize im Eingaberaum wird zufällig ein Reiz v ausgewählt.
- Response: Es wird das Neuron U_c bestimmt, dessen Gewicht dem Reiz am ähnlichsten ist. Ähnlichkeit wird hier durch räumliche Nähe definiert. Dies kann durch Berechnung der euklidischen Distanz zwischen dem Reiz und allen anderen Neuronen(-gewichten) ermittelt werden:

$$|v - w_c| \leq |v - w_r| \forall r \in A$$

- Anschließend werden die Gewichte des Neurons sowie der Neuronen seiner Umgebung angepasst, so dass die Gewichte dem Reiz ähnlicher werden.

$$w_{r-neu} = w_{r-alt} + \eta h(v - w_{r-alt})$$

wobei diese Adaption von der Lernrate und der Umgebung h abhängt.

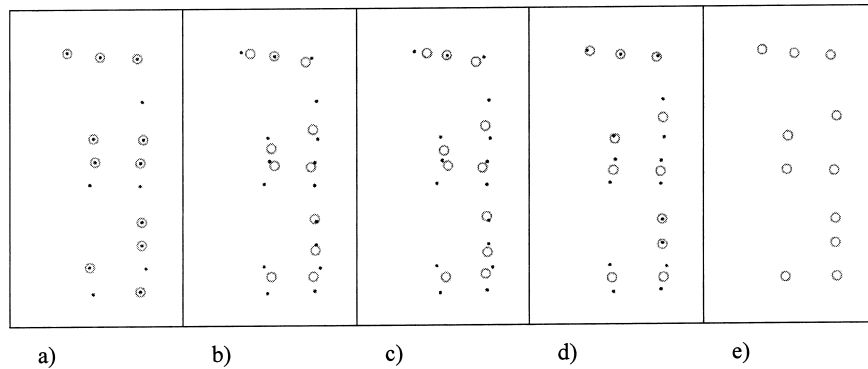


Abb. 6: Typifizierung einer 2D-Struktur durch Reduktion auf 70%: a) Startsituation (dunkle Punkte Stimuli, helle Kreise Neuronen), b)-d) Zwischensituationen der Neuronenverteilung, e) generalisierte Endsituation.

- Diese Lernschritte werden iterativ ausgeführt, bis ein Abbruchkriterium erreicht ist.

Als Funktion h eignet sich die Gaußfunktion, wobei der Radius σ entsprechend der Trainingszeit variiert wird: um die Grobstruktur der Karte zu bilden, sollte er zunächst groß sein, anschließend jedoch kleiner werden, zur Ausbildung der lokalen Feinstruktur. Die Reichweite der Funktion ist damit am Anfang – in der Ordnungsphase – hoch, womit eine optimale Verteilung der Neuronen erreicht wird. Mit wachsender Anzahl an Lernschritten verringert sich die Reichweite und bewirkt, dass die Änderungen nur noch lokal sind, und führt somit zur Konvergenz. Abbildung 6 zeigt die Ausgangssituation, sowie mehrere Zwischenschritte und die generalisierte Endsituation. Deutlich ist die Anpassung der Neuronen (helle Kreise) und die ursprüngliche Situation (dunkle Punkte) zu erkennen.

Durch dieses Wechselspiel zwischen Reizen und Neuronen wird die dichte-basierte Neuverteilung der Objekte erhalten. In der Anwendung der Gebäudegeneralisierung stellen die ursprünglichen Gebäude die Reize dar, die Neuronen bestehen aus der reduzierten Menge der Objekte. In einem iterativen Prozess werden nun die Neuronen an die Reize derart angepasst, dass letztere versuchen, die Neuronen anzuziehen. Dies hat schließlich den Effekt, dass sich an Stellen mit vielen Reizen auch mehr Neuronen anlagern – d.h. die ursprüngliche Situation der Objektverteilung wird nun mit weniger Objekten wiedergegeben.

Abbildung 7 zeigt drei unterschiedliche Generalisierungsstufen eines Datenbestandes.

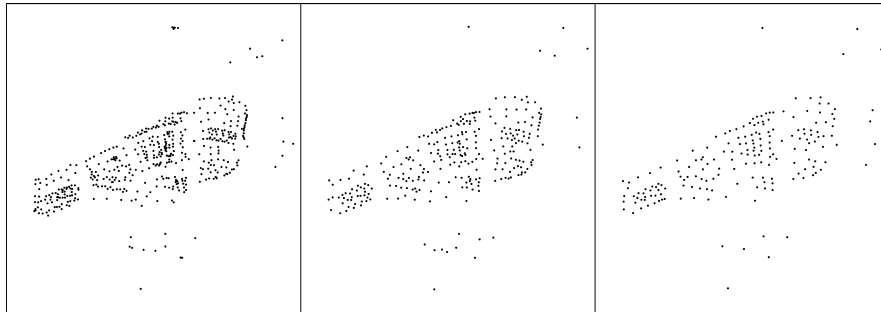


Abb. 7: Ausdünnung eines Datenbestandes (links) mit unterschiedlichen Reduktionsraten: Mitte: 50%, rechts: 30% der ursprünglichen Punkte.



Abb. 8: Generalisierung von Gebäuden für einen Zielmaßstab 1:50.000: links: Überlagerung mit zu detaillierten Gebäuden; rechts: Ergebnis nach Typifizierung.

2.2. Beispiele für die Generalisierung von Gebäuden

Abbildung 8, links, zeigt eine Ausgangssituation in der deutlich wird, dass die Gebäude viel zu klein und zu detailliert sind. Nach der Typifizierung sind die Gebäudesymbole getrennt wahrnehmbar und sie geben in der Summe die korrekte räumliche Verteilung der Gebäude in der Realität wieder.

Abbildung 9 zeigt, wie mit dem Verfahren Gebäudegeneralisierungen für unterschiedliche Zielmaßstäbe automatisch generiert werden können. Die Prozessierung ist jeweils gleich, lediglich sind unterschiedliche Werte als Zielmaßstab anzugeben. In diesem Fall wurden die Maßstäbe 1:25.000, 1:50.000 und 1:75.000 vollautomatisch erzeugt. Solche Verfahren sind nicht nur für die Kartenproduktion relevant, sondern zunehmend auch für die Darstellung räumlicher Information auf kleinen mobilen Displays (SESTER & BRENNER, 2004).

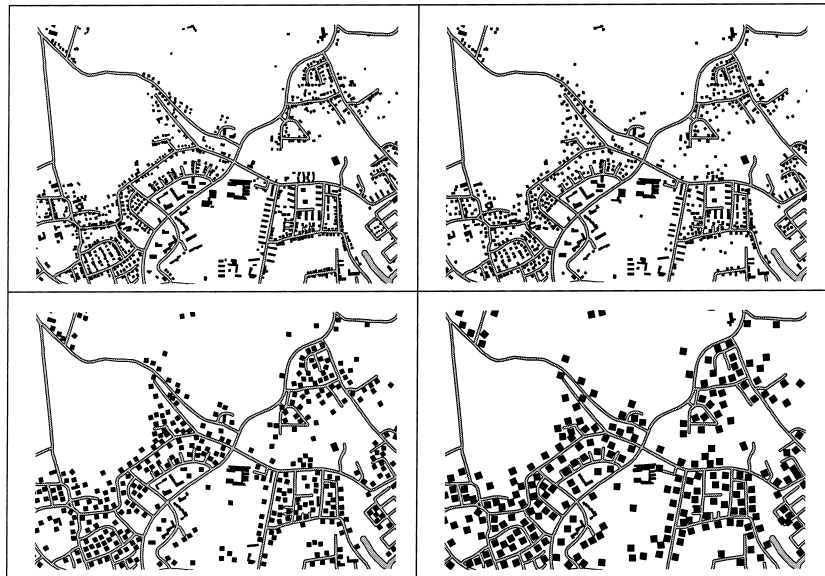


Abbildung 9: Gebäudedarstellung in unterschiedlichen Zielmaßstäben: Ausgangssituation: überlagerte Gebäude Daten (oben links), Zielmaßstäbe 1:25.000 (oben rechts), 1:50.000 (unten links), 1:75.000 (unten rechts).

3. Nutzung von Geodaten für die Navigation

Navigationssysteme sind in Fahrzeugen heute schon fast zur Selbstverständlichkeit geworden. Sollen diese allerdings für die Fußgängernavigation eingesetzt werden, stoßen sie an ihre Grenzen: beispielsweise sind meist Fußgänger-routen nicht im Datensatz vorhanden, etwa Fußgängerzonen, Parkwege oder auch Wege innerhalb von Gebäuden. Ferner sind die Routenanweisungen nicht an die Gewohnheiten und Bedürfnisse des Menschen angepasst. Dieser kann Anweisungen der Form „in 156m rechts abbiegen“ weniger gut verarbeiten als eine, die sogenannte Landmarken nutzt: „biege an der Kirche rechts ab“. B. ELIAS (2005) entwickelte in ihrer Dissertation ein Verfahren, welches ermöglicht, automatisch geeignete Landmarken an Entscheidungspunkten eines Weges zu bestimmen. Als Verfahren wurde eine Methode des Maschinellen Lernens erweitert, so dass sie in der Lage ist, die für Landmarken typischen Eigenschaft, nämlich ihre lokale Einzigartigkeit, automatisch abzuleiten. Dies Landmarken lassen sich direkt als Routenanweisungen verwenden. Weiterhin lassen sie sich als Optimierungskriterium in Routingalgorithmen einbinden, so dass speziell nach „einfachen Routen“ gesucht werden kann, also solchen, die sich besonders gut merken bzw. kommunizieren lassen (ELIAS & SESTER, 2006).

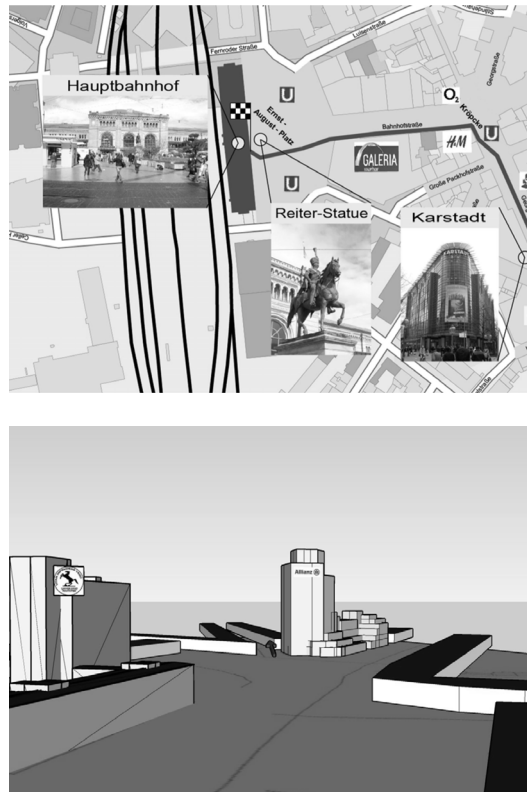


Abb. 10: Routenvisualisierungen mit Landmarken in 2D und in 3D.

Abbildung 10 zeigt ein Beispiel für Visualisierungen von Routen, die durch Landmarken unterstützt sind.

4. Zusammenfassung

Die gezeigten Arbeiten wollen beispielhaft verdeutlichen, dass eine Automation in der raumbezogenen Datenerfassung erforderlich ist, um die anfallenden Massendaten handhaben zu können, um sie sinnvoll auswerten zu können und für neue Anwendungen einzusetzen. Dies erfordert eine Problem- und Nutzer-adaptierte Aufbereitung der Daten. Hierzu werden Methoden der Geodäsie, Informatik und Mathematik eingesetzt. Ziel der Automation ist es, die räumlichen kognitiven Fähigkeiten des Menschen zu verstehen und nachzubilden, um somit Verfahren bereit zu stellen, die ihn zunehmend unterstützen können. Bei-

spiele hierfür sind die Suche nach relevanter Information und die Wegesuche in komplexen Umgebungen.

5. Literatur

- BRENNER, C., 2005: Building reconstruction from images and laser scanning, Int. Journal of Applied Earth Observation and Geoinformation, Theme Issue on „Data Quality in Earth Observation Techniques“, 6(3-4), March 2005, Elsevier, 187-198.
- ELIAS, B., 2006: Extraktion von Landmarken für die Navigation, in: Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover, Heft Nr. 260, DGK, Reihe C, Nr. 596.
- ELIAS, B. & M. SESTER, 2006: Incorporating Landmarks with Quality Measures in Routing Procedures, in: Raubal, M., H.J. Miller, A.U. Frank, M.F. Goodchild (Eds.): Geographic Information Science, GIScience 2006, Münster, Germany, LNCS 4197, pp. 65-80, 2006.
- HEINZLE, F., K.-H. ANDERS & M. SESTER, 2006: Pattern Recognition in Road Networks on the Example of Circular Road Detection, in: Raubal, M., H.J. Miller, A.U. Frank, M.F. Goodchild (Eds.): Geographic Information Science, GIScience 2006, Münster, Germany, LNCS 4197, pp. 253-267, 2006.
- HEINZLE, F., 2007: Verfahren zur Interpretation von Vektordaten, in: Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover, Heft Nr. 267, DGK, Reihe C, Nr. 607.
- KOHONEN, T., 1982, Self-organized formation of topologically correct feature maps. Biological Cybernetics, 42, pp. 59-69.
- SESTER, M., 2005: Optimizing Approaches for Generalization and Data Abstraction, in: International Journal of Geographic Information Science, Vol 19, Nr. 8-9, pp. 871-897, 2005.
- SESTER, M. & C. BRENNER, 2004: Continuous Generalization for Visualization on Small Mobile Devices, in: P. Fisher, Ed., 'Proceedings of the International Symposium on Spatial Data handling (SDH)', Leicester, 2004.